

# Towards a Unified Framework for Approximate Quantum Error Correction

Prabha Mandayam

*The Institute of Mathematical Sciences, Taramani, Chennai - 600113, India.\**

Hui Khoo Ng

*Centre for Quantum Technologies, National University of Singapore,  
Block S15, 3 Science Drive 2, Singapore 117543,*

*Applied Physics Lab, DSO National Laboratories, 20 Science Park Drive, Singapore 118230.\**

(Dated: February 24, 2012)

Operator quantum error correction extends the standard formalism of quantum error correction (QEC) to codes in which only a subsystem within a subspace of states is used to store information in a noise-resilient fashion. Motivated by recent work on approximate QEC, which has opened up the possibility of constructing subspace codes beyond the framework of perfect error correction, we investigate the problem of *approximate* operator quantum error correction (AOQEC). We demonstrate easily checkable sufficient conditions for the existence of AOQEC codes. Furthermore, for certain classes of noise processes, we prove the efficacy of the transpose channel as a simple-to-construct recovery map that works nearly as well as the optimal recovery channel, with optimality defined in terms of worst-case fidelity over all code states. This work generalizes our earlier approach [1] of using the transpose channel for approximate correction of subspace codes to the case of subsystem codes, and brings us closer to a unifying framework for approximate quantum error correction.

The vast majority of existing work on quantum error correction (QEC) focuses on the standard paradigm of perfect error correction [2, 3]. Here, the code  $\mathcal{C}$  and the noise are such that there exists a recovery operation that *completely* removes the effects of the noise on the information stored in the code. Mathematically, this idea is captured by a set of conditions for perfect error correction [4] that must be satisfied by the code as well as the noise process.

That such perfect QEC conditions can be satisfied tends to be special rather than generic. The prototypical example is that of independent noise acting on a few physical qubits, and one finds codes that satisfy the perfect QEC conditions assuming that no more than  $t$  of the qubits have errors. In such a scenario, what is taken as the noise process  $\mathcal{E}$  in the QEC conditions is only the part of the noise that describes the occurrence of  $t$  or fewer errors, while the full physical noise process  $\mathcal{E}_0$  can have more errors, albeit with a lower probability. A code that satisfies the QEC conditions for  $\mathcal{E}$  will thus only satisfy the conditions *approximately* for the full noise process  $\mathcal{E}_0$ . Furthermore, in practice, it is unrealistic to expect full and ideal characterization of the noise process. Thus, a code designed to satisfy the perfect QEC conditions for the *expected* noise process will typically only satisfy those conditions approximately for the true noise process. This motivates the idea of *approximate* quantum error correction (AQEC), where the recovery operation removes most, but not necessarily all, the effects of the noise on the information stored in the code.

Recent studies on AQEC examined the case of *subspace* codes, where information is stored in the entire subspace

corresponding to the code. Examples of AQEC codes that recover the information with fidelity comparable to that of perfect QEC codes, have now been obtained, using both analytical [1, 5–9] and numerical [10–13] approaches. These results suggest that the requirement for perfect recovery may be too stringent for certain tasks and approximate QEC may be more natural and practical.

In [1], we demonstrated a universal, near-optimal recovery map—the *transpose channel* [6, 14]—for AQEC codes. Optimality was defined in terms of the worst-case recovery fidelity over all states in the code. Our analytical approach was a departure from earlier work relying on exhaustive numerical search for the optimal recovery map, with optimality defined based on entanglement fidelity [11, 12, 15]. Furthermore, the analysis in [1] gave a quantitative demonstration of the efficacy of the transpose channel as a good recovery operation, regardless of the noise channel or the code used. In this paper, we extend our approach based on the transpose channel to the more general case of approximate *operator* quantum error correction (AOQEC).

In operator quantum error correction (OQEC), the code has a bipartite tensor-product structure, where one subsystem  $A$  (the *correctable subsystem*) is correctable under action of the noise, while the other subsystem  $B$  (the *noisy subsystem*) can be disturbed by the noise beyond repair [16–18]. The information to be protected against noise is thus stored only in subsystem  $A$ . Subspace codes are a special case of such *subsystem* codes, with a trivial noisy subsystem  $B$ . While this generalization does not lead to new families of codes, it does lead to more efficient decoding procedures [19, 20], and hence to better fault-tolerant schemes and improved bounds on the accuracy threshold [21]. Starting with the Bacon-Shor codes—a family of subsystem codes arising from

---

\*Both authors contributed equally to this work.

Shor's 9-qubit code [20]—several examples of perfectly correctable stabilizer subsystem codes have been constructed [22].

In this paper, we begin by proving a set of perfect OQEC conditions in Section II that is completely equivalent to the perfect OQEC conditions found in [16, 18]. This alternate set of conditions makes the role of the transpose channel in perfect OQEC manifestly clear. Furthermore, it serves as a natural starting point for perturbation to a set of sufficient conditions for approximate OQEC (Section III A). We then proceed to show the near-optimality of the transpose channel recovery map for AOQEC under three different scenarios: (i) The noisy subsystem  $B$  starts in a maximally mixed state (Section III B 1); (ii) Subsystem  $B$  is in fact perfectly correctable (Section III B 2); (iii) The noise process destroys nearly all information that may be present in the noisy subsystem (Section III B 3).

## I. BASIC DEFINITIONS

We consider a decomposition of the Hilbert space of our quantum system

$$\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B + K. \quad (1)$$

Suppose we wish to store information in the  $\mathcal{H}_A$  factor. We treat  $\mathcal{H}_{AB} \equiv \mathcal{H}_A \otimes \mathcal{H}_B$  as if it is the Hilbert space of a composite system comprising two subsystems  $A$  and  $B$ . In reality, subsystems  $A$  and  $B$  may not correspond to “natural” separate physical degrees of freedom of the system, but as only as mathematical factors in a decomposition as in Eq. (1). We denote the projector onto  $\mathcal{H}_{AB}$  as  $P$ .  $P$  can also be written as a tensor-product:  $P = P_A \otimes P_B$ , where  $P_{A(B)}$  is the projector onto  $\mathcal{H}_{A(B)}$ .

Information is stored as a choice between the states of subsystem  $A$ . The state on subsystem  $B$  can be arbitrary, since it carries no information. More concretely, we make use of a *code*  $\mathcal{C}$ , comprising all product states on  $AB$  [28]:

$$\mathcal{C} \equiv \{\rho = \rho_A \otimes \rho_B \mid \forall \rho_A \in \mathcal{S}(\mathcal{H}_A), \rho_B \in \mathcal{S}(\mathcal{H}_B)\}, \quad (2)$$

where  $\mathcal{S}(\mathcal{H}_{A(B)})$  denotes the set of all states (density operators) on subsystem  $A(B)$ . The information is stored only in subsystem  $A$  in that two states  $\rho_A \otimes \tau_B$  and  $\rho_A \otimes \sigma_B$  differing only in the state of  $B$  corresponds to the same encoded information. Standard quantum error correction (QEC) deals with codes where subsystem  $B$  is of trivial dimension; operator quantum error correction (OQEC) is the generalization into situations where  $B$  can be nontrivial.

We wish to examine the longevity of the information stored in subsystem  $A$  in the presence of noise [29]. We describe the noise by a quantum channel or map acting on  $AB$ , i.e., a completely positive (CP), trace-preserving (TP) map  $\mathcal{E} : \mathcal{B}(\mathcal{H}_{AB}) \rightarrow \mathcal{B}(\mathcal{P}_{\mathcal{E}})$ . Here, the notation  $\mathcal{B}(\mathcal{V})$  refers to the set of all bounded operators on a vector space  $\mathcal{V}$ .  $\mathcal{P}_{\mathcal{E}}$  is the support of  $\mathcal{E}(\mathcal{B}(\mathcal{H}_{AB}))$ , or equivalently

the support of  $\mathcal{E}(P)$ .  $\mathcal{E}$  can be specified via a set of Kraus operators  $\{E_i\}_{i=1}^N$ , so that  $\mathcal{E}$  acts as

$$\mathcal{E}(\rho) = \sum_{i=1}^N E_i \rho E_i^\dagger. \quad (3)$$

That  $\mathcal{E}$  is TP translates into the statement  $\sum_i E_i^\dagger E_i = P$ . The Kraus representation of a CPTP channel is non-unique: if  $\{E_i\}$  is a Kraus representation of  $\mathcal{E}$ , then  $\{F_j \equiv \sum_i u_{ij} E_i\}$  is a Kraus representation of the same channel  $\mathcal{E}$ , where  $u_{ij}$  are matrix elements of a unitary matrix. A recovery operation  $\mathcal{R} : \mathcal{B}(\mathcal{P}_{\mathcal{E}}) \rightarrow \mathcal{B}(\mathcal{H}_{AB})$  performed on the code, after each application of the noise  $\mathcal{E}$ , to attempt to reverse the effects of the noise is also described as a CPTP map.

Since information is stored in subsystem  $A$  only, we are concerned only with how well the noise preserves the information initially stored in  $A$ , while any state on  $B$  can be distorted beyond repair by the noise. Heuristically, we say that a code  $\mathcal{C}$  is *approximately correctable* under noise  $\mathcal{E}$  if and only if there exists a CPTP recovery map  $\mathcal{R}$  such that

$$\text{tr}_B[(\mathcal{R} \circ \mathcal{E})(\rho)] \simeq \text{tr}_B(\rho) \quad \forall \rho \in \mathcal{C}, \quad (4)$$

where  $\text{tr}_B(\cdot)$  denotes the partial trace over subsystem  $B$ .

This heuristic notion can be formalized by quantifying the deviation of the recovered state from the initial encoded state by the fidelity between the two states. The fidelity between two states  $\rho$  and  $\sigma$  is  $F(\rho, \sigma) \equiv \text{tr} \sqrt{\rho^{1/2} \sigma \rho^{1/2}}$ , which for a pure state  $\rho$ , can be written as

$$F(|\psi\rangle, \sigma) \equiv \sqrt{\langle \psi | \sigma | \psi \rangle}. \quad (5)$$

We define as the *fidelity loss for state*  $\rho$ ,  $\eta_{\mathcal{R}}\{\rho\}$ , under noise  $\mathcal{E}$  and recovery  $\mathcal{R}$ , as the deviation from 1 of the square of the fidelity between the initial state  $\rho$  and the recovered state  $(\mathcal{R} \circ \mathcal{E})(\rho)$ , i.e.,

$$\eta_{\mathcal{R}}\{\rho\} \equiv 1 - F^2(\text{tr}_B(\rho), \text{tr}_B[(\mathcal{R} \circ \mathcal{E})(\rho)]). \quad (6)$$

The performance of a recovery  $\mathcal{R}$  on a code  $\mathcal{C}$  is then characterized by the *fidelity loss for*  $\mathcal{C}$  given by

$$\eta_{\mathcal{R}}\{\mathcal{C}\} \equiv \max_{\rho \in \mathcal{C}} \eta_{\mathcal{R}}\{\rho\}. \quad (7)$$

How well  $\mathcal{R}$  recovers the information initially stored in subsystem  $A$  is hence gauged by the *worst-case fidelity* (over all states in the code) between the initial and recovered states. Because the fidelity is jointly concave in its arguments, the worst-case fidelity is always attained on a pure state on  $AB$ . The maximization in Eq. (7) can thus be restricted to pure states on  $AB$  only. Often, when the meaning is clear from the context, we will drop the argument from  $\eta_{\mathcal{R}}\{\mathcal{C}\}$  and simply write  $\eta_{\mathcal{R}}$ .

Let  $\mathcal{R}_{\text{op}}$  be the recovery map with the smallest fidelity loss among all possible recovery maps for code  $\mathcal{C}$ , i.e.,

$$\eta_{\text{op}}\{\mathcal{C}\} \equiv \eta_{\mathcal{R}_{\text{op}}}\{\mathcal{C}\} = \min_{\mathcal{R}} \eta_{\mathcal{R}}\{\mathcal{C}\}. \quad (8)$$

We refer to  $\mathcal{R}_{\text{op}}$  as the *optimal recovery*, and  $\eta_{\text{op}}$  as the *optimal fidelity loss*. As is clear from the definition, whether or not a recovery map is optimal depends on the code in question.

A code  $\mathcal{C}$  with  $\eta_{\text{op}} = 0$  under noise  $\mathcal{E}$  is said to be *perfectly correctable* on  $A$  under  $\mathcal{E}$ . In general, we say that a code is  $\epsilon$ -*correctable* on  $A$  under noise  $\mathcal{E}$  if  $\epsilon \geq \eta_{\text{op}}$ , which means that it is possible to recover the information stored in  $A$  with a worst-case fidelity no smaller than  $\sqrt{1 - \epsilon}$ . Sometimes, to compare with our previous work on subspace codes, we will refer to the scenario with nontrivial  $B$  subsystem as approximate *operator* quantum error correction (AOQEC); when  $B$  is trivial, i.e., subspace codes, we refer to this case as approximate quantum error correction (AQEC).

Central to our analysis is a recovery map built from the noise channel and code, known as the *transpose channel*. The transpose channel, denoted as  $\mathcal{R}_P$ , is defined, in a manifestly representation-independent way, as

$$\mathcal{R}_P \equiv \mathcal{P}_{\mathcal{C}} \circ \mathcal{E}^\dagger \circ \mathcal{N}. \quad (9)$$

Here,  $\mathcal{E}^\dagger$  is the adjoint of  $\mathcal{E}$ , i.e., the channel with Kraus operators  $\{E_i^\dagger\}_{i=1}^N$  if  $\mathcal{E}$  has Kraus operators  $\{E_i\}_{i=1}^N$ .  $\mathcal{N}$  is a normalization map  $\mathcal{N}(\cdot) \equiv \mathcal{E}(P)^{-1/2}(\cdot)\mathcal{E}(P)^{-1/2}$  (the inverse is taken on the support of  $\mathcal{E}(P)$ ).  $\mathcal{P}_{\mathcal{C}}$  is the projection onto the support of  $\mathcal{C}$ , i.e.,  $\mathcal{P}_{\mathcal{C}}(\cdot) = P(\cdot)P$ . One can write  $\mathcal{R}_P$  explicitly in terms of its Kraus operators  $\{R_i^P\}_{i=1}^N$ , where

$$R_i^P = PE_i^\dagger \mathcal{E}(P)^{-1/2}. \quad (10)$$

$\mathcal{R}_P$  is trace-preserving (TP) on  $\mathcal{P}_{\mathcal{E}}$ . We denote the fidelity loss for using the transpose channel as the recovery by  $\eta_P$ .

The term “transpose channel” owes its origin to [14], where this channel was first defined in an information-theoretic context. It was shown [23] that the transpose channel has the property of being the unique noise channel that saturates Uhlmann’s theorem on the monotonicity of relative entropy—a fact that was later used to characterize states that saturate the strong subadditivity of quantum entropy [24].  $\mathcal{R}_P$  is a special case of a recovery map introduced in [6] for reversing the effects of a quantum channel on a given initial state. In fact, in that context,  $\mathcal{R}_P$  is exactly the recovery map for the initial state  $P/d$ , where  $d$  is the dimension of  $\mathcal{C}$ . In [25, 26],  $\mathcal{R}_P$  was shown to be useful for correcting information carried by codes preserved according to an operationally motivated notion.

## II. PERFECT OQEC: TWO EQUIVALENT CONDITIONS

We begin our study with the case of perfect OQEC, where there exists a recovery map such that the fidelity of any state on  $A$  after noise and recovery attains the maximal value of 1. Necessary and sufficient algebraic

conditions for the existence of an OQEC code for a given channel  $\mathcal{E} \sim \{E_i\}$  were found in [16–18]. Here, we prove an equivalent set of OQEC conditions that is more useful than the original OQEC conditions from the perspective of approximate error correction.

**Theorem 1.** *Given a CP channel  $\mathcal{E} : \mathcal{B}(\mathcal{H}_{AB}) \rightarrow \mathcal{B}(\mathcal{P}_{\mathcal{E}})$  with a set of Kraus operators  $\{E_i\}$ , the following two statements are equivalent:*

- (A)  $PE_i^\dagger \mathcal{E}(P)^{-1/2} E_j P = P_A \otimes B_{ij}$ , for all  $i, j$ , and  $B_{ij} \in \mathcal{B}(\mathcal{H}_B)$ ;
- (B)  $PE_i^\dagger E_j P = P_A \otimes B'_{ij}$ , for all  $i, j$ , and  $B'_{ij} \in \mathcal{B}(\mathcal{H}_B)$ .

Statement (B) is the form of the OQEC conditions given in [16–18]—a subsystem code  $\mathcal{C}$  built on composite system  $AB$  is perfectly correctable on subsystem  $A$  under the action of  $\mathcal{E}$  if and only if Statement (B) holds.

Before proving Theorem 1, let us first point out the usefulness of this new form of the OQEC conditions (Statement (A)). Observe that the left-hand-side of the equation contained in Statement (A) is nothing but a Kraus operator  $R_i^P E_j$  of the channel  $\mathcal{R}_P \circ \mathcal{E}$ . Statement (A) thus says that the transpose channel  $\mathcal{R}_P$  is the recovery map needed to recover perfectly the state on subsystem  $A$  after the action of the channel  $\mathcal{E}$ . The correctability of codes satisfying the error correction conditions become manifestly clear with this new form of the conditions. A special case of Theorem 1 was derived in [1] for subspace codes.

Now, let us prove Theorem 1:

*Proof.* (A) $\Rightarrow$ (B): For any  $i, j$ ,  $\sum_k (P_A \otimes B_{ik})(P_A \otimes B_{kj}) = \sum_k (PE_i^\dagger \mathcal{E}(P)^{-1/2} E_k P)(PE_k^\dagger \mathcal{E}(P)^{-1/2} E_j P) = PE_i^\dagger E_j P$ , which gives  $PE_i^\dagger E_j P = P_A \otimes B'_{ij}$ , with  $B'_{ij} \equiv \sum_k B_{ik} B_{kj}$ .

(B) $\Rightarrow$ (A): Let  $\{|s\rangle_B\}$  be an orthonormal basis for  $\mathcal{H}_B$ . Statement (B) implies

$$P_A E_{is}^\dagger E_{jt} P_A = \lambda_{(is)(jt)} P_A, \quad (11)$$

where  $E_{is} \equiv E_i |s\rangle_B$  is an operator that brings vectors in  $\mathcal{H}_A$  to vectors in  $\mathcal{H}_{AB}$ , and  $\lambda_{(is)(jt)} \equiv \langle s | B'_{ij} | t \rangle$ .  $\{E_{is}\}$  is a set of Kraus operators for the CP channel  $\mathcal{E}_A : \mathcal{B}(\mathcal{H}_A) \rightarrow \mathcal{B}(\mathcal{P}_{\mathcal{E}})$  defined by  $\mathcal{E}_A(\rho_A) = \mathcal{E}(\rho_A \otimes P_B)$ . We view  $\Lambda \equiv (\lambda_{(is)(jt)})$  as a two-index matrix, where the first index is the double index  $(is)$ , and the second is  $(jt)$ . Observe that  $\lambda_{(jt)(is)}^* = \lambda_{(is)(jt)}$ , i.e.,  $\Lambda$  is a hermitian matrix. It is thus diagonalizable, i.e.,  $\exists U \equiv (u_{(is)(jt)})$  such that  $U \Lambda U^\dagger = \Lambda_D$ , where  $\Lambda_D$  is a diagonal matrix. More explicitly, we have

$$\sum_{(i's'), (j't')} u_{(is)(i's')} \lambda_{(i's')(j't')} u_{(jt)(j't')}^* = \delta_{(is)(jt)} d_{is}, \quad (12)$$

where  $d_{is}$  are the diagonal entries of  $\Lambda_D$ . Using this, we can write Eq. (11) in its diagonal form:

$$P_A F_{is}^\dagger F_{jt} P_A = \delta_{ij} \delta_{st} d_{is} P_A, \quad (13)$$

where  $F_{is} \equiv \sum_{(i's')} u_{(is)(i's')}^* E_{i's'}$  gives a different Kraus representation for  $\mathcal{E}_A$ . Equation (13) gives the polar decomposition  $F_{is} P_A = \sqrt{d_{is}} V_{is} P_A$ , where  $V_{is}$  is a unitary operator satisfying  $P_A V_{is}^\dagger V_{jt} P_A = \delta_{ij} \delta_{st} P_A$ . Let  $P_{is} \equiv V_{is} P_A V_{is}^\dagger$ . Then,  $P_{is}$ 's are orthogonal projectors, since  $P_{is} P_{jt} = \delta_{ij} \delta_{st} P_{is}$ . Direct computation gives  $\mathcal{E}(P) = \sum_{is} d_{is} P_{is}$ , i.e.,  $\mathcal{E}(P)$  is a sum of orthogonal projectors, and hence easy to invert:  $\mathcal{E}(P)^{-1/2} = \sum_{is} d_{is}^{-1/2} P_{is}$ . Further algebra gives  $P E_i^\dagger \mathcal{E}(P)^{-1/2} E_j P = P_A \otimes B_{ij}$ , with  $B_{ij} \equiv \sum_{st} \sum_{kv} u_{(kv)(is)}^* u_{(kv)(jt)} \sqrt{d_{kv}} |s\rangle_B \langle t|$ .  $\square$

### III. APPROXIMATE OQEC CONDITIONS

#### A. Sufficient AOQEC conditions

For a TP  $\mathcal{E}$ , our perfect OQEC conditions (Statement A of Theorem 1) state that  $\mathcal{R}_P \circ \mathcal{E}$  acts as the identity channel on subsystem  $A$ . Perturbing Statement A by adding to the right-hand-side of the equation a small correction to  $P_A \otimes B_{ij}$  will modify the previous sentence to say that  $\mathcal{R}_P \circ \mathcal{E}$  acts *nearly* as the identity channel on subsystem  $A$ . This provides a natural route to sufficient conditions for AOQEC: If the perturbation to Statement A is small enough, the code is  $\epsilon$ -correctable on  $A$  with  $\epsilon$  small. What remains is to relate quantitatively the size of the perturbation to  $\epsilon$ .

**Theorem 2.** *Consider a CPTP noise channel  $\mathcal{E} \sim \{E_i\}$  and a code  $\mathcal{C}$  on  $\mathcal{H}_{AB}$  as defined above. Suppose*

$$P E_i^\dagger \mathcal{E}(P)^{-1/2} E_j P = P_A \otimes B_{ij} + \Delta_{ij}, \quad \forall i, j, \quad (14)$$

for  $B_{ij} \in \mathcal{B}(\mathcal{H}_B)$ , and  $\Delta_{ij} \in \mathcal{B}(\mathcal{H}_{AB})$ . Then,  $\mathcal{C}$  is  $\epsilon$ -correctable on  $A$  under  $\mathcal{E}$  for  $\epsilon \geq \eta_P$ , where

$$\eta_P \equiv \max_{|\psi_A, \phi_B\rangle} \langle \phi_B | \sum_{ij} \left[ \langle \psi_A | \Delta_{ij}^\dagger \Delta_{ij} | \psi_A \rangle - \langle \psi_A | \Delta_{ij}^\dagger | \psi_A \rangle \langle \psi_A | \Delta_{ij} | \psi_A \rangle \right] | \phi_B \rangle. \quad (15)$$

*Proof.* The TP condition on  $\mathcal{R}_P \circ \mathcal{E}$  gives the relation  $P = \sum_{ij} [P_A \otimes B_{ij}^\dagger B_{ij} + \Delta_{ij}^\dagger \Delta_{ij} + (P_A \otimes B_{ij}^\dagger) \Delta_{ij} + \Delta_{ij}^\dagger (P_A \otimes B_{ij})]$ . Using this, direct computation gives

$$\begin{aligned} F^2[|\psi_A\rangle, (\text{tr}_B \circ \mathcal{R}_P \circ \mathcal{E})(|\psi_A, \phi_B\rangle \langle \psi_A, \phi_B|)] \\ = 1 - \langle \psi_A, \phi_B | \sum_{ij} \Delta_{ij}^\dagger (P_A - |\psi_A\rangle \langle \psi_A|) \otimes P_B \Delta_{ij} | \psi_A, \phi_B \rangle. \end{aligned} \quad (16)$$

This yields the expression for  $\eta_P$  in Eq. (15) upon recalling that the worst-case fidelity is attained on a pure state on  $AB$ .  $\square$

While the bound for  $\epsilon$  in Theorem 2 is tight, the maximization over all pure product states on  $AB$  in the expression for  $\eta_P$  may not be easy to evaluate. Instead, we can relax the bound and obtain a simpler (but weaker) sufficiency condition:

**Corollary 3.** *Given the same conditions as in Theorem 2,  $\mathcal{C}$  is  $\epsilon$ -correctable on  $A$  under  $\mathcal{E}$  if*

$$\epsilon \geq \left\| \sum_{ij} \Delta_{ij}^\dagger \Delta_{ij} \right\|, \quad (17)$$

where  $\|\cdot\|$  is the operator norm.

*Proof.* Observe that, for any pure product state  $|\psi_A, \phi_B\rangle$ , the expression in Eq. (15) to be maximized is bounded from above by  $\langle \psi_A, \phi_B | \sum_{ij} \Delta_{ij}^\dagger \Delta_{ij} | \psi_A, \phi_B \rangle \leq \left\| \sum_{ij} \Delta_{ij}^\dagger \Delta_{ij} \right\|$ . This gives  $\eta_P \leq \left\| \sum_{ij} \Delta_{ij}^\dagger \Delta_{ij} \right\|$ , which immediately yields the corollary statement.  $\square$

Note that, in both Theorem 2 and Corollary 3, we obtain the analogous subspace code results found in [1] when  $B$  is restricted to a trivial subsystem.

#### B. Towards necessary AOQEC conditions

In [1], the transpose channel was shown to be near-optimal for subspace codes  $\mathcal{C}$ , i.e., its fidelity loss for code  $\mathcal{C}$  under noise  $\mathcal{E}$  is close to the optimal fidelity loss. Here, we repeat the quantitative statement of the near-optimality of the transpose channel [1], adapted to the language suited for this paper:

**Theorem 4** (Corollary 4 of [1]). *Consider a subspace code  $\mathcal{C}$  ( $B$  is trivial), with  $d_A$  denoting the dimension of  $\mathcal{H}_A$ , and optimal fidelity loss  $\eta_{op}$  under CPTP noise channel  $\mathcal{E}$ . Then, the fidelity loss  $\eta_P$  for the transpose channel satisfies*

$$\eta_{op} \leq \eta_P \leq \eta_{op} f(\eta_{op}; d_A), \quad (18)$$

where  $f(\eta; d)$  is the function

$$f(\eta; d) \equiv \frac{(d+1) - \eta}{1 + (d-1)\eta} = (d+1) + O(\eta). \quad (19)$$

The left inequality  $\eta_{op} \leq \eta_P$  of Eq. (18) is true simply by definition of  $\eta_{op}$ . The proof of the right inequality  $\eta_P \leq \eta_{op} f(\eta_{op}; d_A)$  in [1] requires the following inequality that holds for any pure state  $\psi_A \equiv |\psi_A\rangle \langle \psi_A|$  from a subspace code  $\mathcal{C}$ :

$$1 - \eta_{op}\{\psi_A\} \leq \sqrt{[1 + (d_A - 1)\eta_{op}\{\mathcal{C}\}][1 - \eta_P\{\psi_A\}]}. \quad (20)$$

Inverting Eq. (20) and recalling the definitions of  $\eta_{op}$  and  $\eta_P$  as the maximization of  $\eta_{(\cdot)}\{\psi_A\}$  over all states in the code yields the right inequality of Eq. (18).

Equation (18) implies that an approximately correctable subspace code must necessarily be such that the fidelity loss for the transpose channel is small. A small fidelity loss for the transpose channel requires that  $\mathcal{E}$  has Kraus operators that satisfy Eq. (14) with  $\Delta_{ij}$  small [30]. This means that Eq. (14) with  $\Delta_{ij}$  small is not only sufficient (as shown in Sec. III A), but also necessary for subspace codes, as was pointed out in [1].



Similarly, generalizing Eq. (18) to subsystem codes where  $B$  is nontrivial will lead to necessary conditions for AOQEC of the same form as the sufficient conditions given in Eq. (14). Unfortunately, we have not been able to extend the right inequality of Eq. (18) to the subsystem case in general. We believe that our inability to prove a similar result when  $B$  is nontrivial lies not with the failure of the transpose channel as a good recovery, but in our proof technique. Alternative approaches are currently being explored.

Despite our current difficulties with the general case, statements similar to Eq. (18) do hold for restricted classes of subsystem codes and channels, and the transpose channel is provably near-optimal for these cases. An obvious case where this works is, of course, one where  $\mathcal{E}$  is a product channel, i.e.,  $\mathcal{E}(\rho_A \otimes \rho_B) = \mathcal{F}_A(\rho_A) \otimes \mathcal{F}_B(\rho_B)$ . For such a channel  $\mathcal{E}$ , the transpose channel is also a product channel, namely, the product of the respective transpose channels of  $\mathcal{F}_A$  and  $\mathcal{F}_B$ . Since there is no flow of information between  $A$  and  $B$ , whether subsystem  $A$  is correctable relies only on the properties of  $\mathcal{F}_A$ . We can thus treat this case as if it is a subspace code on  $A$  under noise  $\mathcal{F}_A$ , for which the transpose channel is indeed near-optimal from Theorem 4. In the remainder of the paper, we discuss three other scenarios where the transpose channel is also near-optimal.

### 1. Maximally mixed state on subsystem $B$

Suppose we do not allow arbitrary states on  $B$ , but restrict our code to comprise only states with the maximally mixed state on  $B$ . We consider codes of the form

$$\mathcal{C}_0 \equiv \left\{ \rho_A \otimes \frac{P_B}{d_B}, \quad \forall \rho_A \in \mathcal{S}(\mathcal{H}_A) \right\}. \quad (21)$$

Here,  $d_B$  is the dimension of  $\mathcal{H}_B$ . Such a code is of practical relevance whenever one lacks control over subsystem  $B$ . Full control over subsystem  $A$  alone is sufficient to guarantee preparation of a product code state. Subsystem  $B$  can be in a random state, which will be well described by the maximally mixed state.

For states in  $\mathcal{C}_0$ , we can write the action of the noise channel  $\mathcal{E}$  as

$$\mathcal{E}\left(\rho_A \otimes \frac{P_B}{d_B}\right) = \sum_{is} \bar{E}_{is} \rho_A \bar{E}_{is}^\dagger \equiv \bar{\mathcal{E}}_A(\rho_A). \quad (22)$$

$\bar{\mathcal{E}}_A$  is a CPTP channel on  $A$  with Kraus operators  $\{\bar{E}_{is} \equiv (1/\sqrt{d_B})E_i|s_B\rangle\}$ , where  $\{|s_B\rangle\}_{s=1}^{d_B}$  is an orthonormal basis for  $\mathcal{H}_B$ . We can then forget about subsystem  $B$  and ask about correctability of code  $\mathcal{C}_0$ —now viewed as a subspace code on  $A$ —under the noise  $\bar{\mathcal{E}}_A$ . Theorem 4 applies and ensures that the transpose channel of  $\bar{\mathcal{E}}_A$ , which we denote as  $\mathcal{R}_{A,P}$ , has fidelity loss close to that of the optimal recovery  $\mathcal{R}_{A,op}$ .

Observe that, for any state in  $\mathcal{C}_0$ , the action of  $\text{tr}_B \circ \mathcal{R}_P \circ \mathcal{E}$  is equal to applying the map  $\mathcal{R}_{A,P} \circ \bar{\mathcal{E}}_A$ , where  $\mathcal{R}_{A,P}$  is the transpose channel for  $\bar{\mathcal{E}}_A$ . Furthermore, the optimal recovery  $\mathcal{R}_{op}$  for the noise  $\mathcal{E}$  acting on  $\mathcal{C}_0$  is such that  $\text{tr}_B \circ \mathcal{R}_{op} = \mathcal{R}_{A,op}$  on input from  $\mathcal{E}(\mathcal{C}_0)$ . Near-optimality of the transpose channel of  $\bar{\mathcal{E}}_A$  thus immediately translates into near-optimality of the transpose channel of  $\mathcal{E}$  for the code  $\mathcal{C}_0$ . We gather these observations into a corollary:

**Corollary 5.** *Consider the code  $\mathcal{C}_0$  defined in Eq. (21) under noise  $\mathcal{E}$ . The transpose channel is near-optimal in that*

$$\eta_P\{\mathcal{C}_0\} \leq \eta_{op}\{\mathcal{C}_0\} f(\eta_{op}\{\mathcal{C}_0\}; d_A) \quad (23)$$

for  $f(\eta; d)$  defined in Eq. (19).

### 2. $B$ is perfectly correctable

Suppose subsystem  $B$  is in fact perfectly correctable, but we choose to use subsystem  $A$  to store the information. This is relevant, for example, when  $B$  corresponds to a degree of freedom that is experimentally inaccessible or uncontrollable, or if  $A$  is a much larger space with greater storage capacity than  $B$ . The transpose channel is again near-optimal in this case.

We begin by showing that, for  $B$  perfectly correctable, the fidelity for a pure initial state on subsystem  $A$  from using the transpose channel as recovery is independent of the initial state of subsystem  $B$ .

**Lemma 6.** *If subsystem  $B$  is perfectly correctable under noise  $\mathcal{E}$ , then  $F[|\psi\rangle_A, (\text{tr}_B \circ \mathcal{R}_P \circ \mathcal{E})(\psi_A \otimes \rho_B)]$ , where  $\psi_A \equiv |\psi_A\rangle\langle\psi_A|$ , is independent of  $\rho_B$ .*

*Proof.*  $B$  perfectly correctable under noise  $\mathcal{E}$  implies the perfect OQEC conditions (Statement (A) of Theorem 1 with the roles of  $A$  and  $B$  interchanged): there exists operators  $A_{ij}$  on  $A$  for all  $i, j$  such that  $PE_i^\dagger \mathcal{E}(P)^{-1/2} E_j P = A_{ij} \otimes P_B$ . Using this, we have

$$\begin{aligned} & F^2[|\psi_A\rangle, (\text{tr}_B \circ \mathcal{R}_P \circ \mathcal{E})(\psi_A \otimes \rho_B)] \\ &= \sum_{ij} \langle \psi_A | \text{tr}_B [A_{ij} \otimes P_B (\psi_A \otimes \rho_B) A_{ij}^\dagger \otimes P_B] | \psi_A \rangle \\ &= \sum_{ij} |\langle \psi_A | A_{ij} | \psi_A \rangle|^2, \end{aligned} \quad (24)$$

which is independent of  $\rho_B$ .  $\square$

Lemma 6 implies the following sequence of inequalities (we remind the reader that  $\mathcal{C}$  here is the code defined in

Eq. (2), where the state on  $B$  can be arbitrary):

$$\begin{aligned}
\eta_P\{\mathcal{C}\} &= \max_{\rho \in \mathcal{C}} \eta_P\{\rho\} \leq \max_{\rho=|\psi_A\rangle\langle\psi_A| \otimes \rho_B} \eta_P\{\rho\} \\
&= \max_{|\psi\rangle_A} \eta_P\{\psi_A \otimes P_B/d_B\} \\
&= \eta_P\{\mathcal{C}_0\} \\
&\leq \eta_{\text{op}}\{\mathcal{C}_0\} f(\eta_{\text{op}}\{\mathcal{C}_0\}; d_A) \\
&\leq \eta_{\text{op}}\{\mathcal{C}\} f(\eta_{\text{op}}\{\mathcal{C}\}; d_A). \quad (25)
\end{aligned}$$

The first inequality follows from the concavity of the fidelity. The second line makes use of Lemma 6. In the last inequality, we have used the fact that  $\mathcal{C}_0 \subset \mathcal{C}$  so that  $\eta_{\text{op}}\{\mathcal{C}\} \geq \eta_{\text{op}}\{\mathcal{C}_0\}$ , and that  $\eta f(\eta; d)$  is a monotonically increasing function of  $\eta$ . Equation (25) gives exactly the right inequality in Eq. (18) applied to the current scenario, from which we draw the conclusion that the transpose channel is also near-optimal on  $A$  under channel  $\mathcal{E}$  with  $B$  perfectly correctable.

### 3. $\mathcal{E}$ destroys information on $B$

Suppose the noise  $\mathcal{E}$  satisfies the following condition:

**Condition 1.** For CPTP  $\mathcal{E}$ , suppose there exists  $\delta \geq 0$  such that

$$\left\| \mathcal{E}(\rho_A \otimes \rho_B) - \mathcal{E}\left(\rho_A \otimes \frac{P_B}{d_B}\right) \right\|_{tr} \leq \delta \left\| \rho_B - \frac{P_B}{d_B} \right\|_{tr} \quad (26)$$

for all states  $\rho_A \in \mathcal{S}(\mathcal{H}_A)$  and  $\rho_B \in \mathcal{S}(\mathcal{H}_B)$ .  $\|O\|_{tr}$  denotes the trace norm of  $O$  given by  $\text{tr}|O|$ .

If  $\delta \ll 1$ , any two states on  $\mathcal{H}_B$ , after the action of  $\mathcal{E}$ , become close together and nearly indistinguishable (as quantified by the trace norm used in Condition 1), corresponding to loss of any information stored as states in  $B$ . A very simple example is a channel  $\mathcal{E}$  that maps all states on  $\mathcal{H}_B$  to some fixed state  $\tau_B$ , for which  $\delta$  can be chosen to be zero. While we have chosen, for convenience of the subsequent analysis, to state Condition 1 in terms of comparing states on  $B$  before and after the channel  $\mathcal{E}$  to what happens to the maximally mixed state  $P_B/d_B$ , one is free to choose other reference states on  $B$  if desired.

For channels and codes satisfying Condition 1, the transpose channel also works well as a recovery channel, as encapsulated in the following corollary:

**Corollary 7.** Given that Condition 1 is satisfied, for a subsystem code  $\mathcal{C}$ ,

$$\eta_P \leq (d_A + 1)\eta_{\text{op}} + 3\delta + O(\delta^2, \eta_{\text{op}}^2, \delta\eta_{\text{op}}). \quad (27)$$

The proof of this corollary is detailed in Appendix A. The idea behind the proof is to first show that the transpose channel works well, compared to the optimal recovery for  $\mathcal{C}$ , as a recovery for the information stored in  $A$  when  $B$  is initially in the maximally mixed state. Since Condition 1 says that  $\mathcal{E}$  brings code states with different states

on  $B$  close together, if the transpose channel works well as a recovery for  $B$  being initially in the maximally mixed state, it will also work well when  $B$  is initially in a different state.

Corollary 7, like similar statements before, tells us that the fidelity loss corresponding to the transpose channel is not much worse than that of the optimal recovery. The additional fidelity loss suffered from using the simpler transpose channel is governed by  $d_A$ , the dimension of the information-carrying subsystem, as well as the parameter  $\delta$  which characterizes how badly  $\mathcal{E}$  destroys distinguishability between states on subsystem  $B$ .

## IV. CONCLUSION

We have studied the role of the transpose channel in operator quantum error correction. We first generalize our alternate form of the perfect QEC conditions involving the transpose channel, to the case of OQEC codes. This completes our understanding as to why certain channels admit perfectly correctable codes, in a particularly intuitive way. Our perfect OQEC conditions naturally lead to sufficient conditions for AOQEC, since the resilience to noise of the information stored in the code can now be quantified in a particularly easy way. We have also demonstrated that the transpose channel works nearly as well as any other recovery channel for three different scenarios of codes and noise. In all three cases, the near-optimality of the transpose channel relies only on  $d_A$ , the dimension of the information-carrying subsystem  $A$ , and not on  $d_B$ , the dimension of the noisy subsystem that carries no information.

Proving necessary conditions for AOQEC based on our transpose channel approach will provide the final missing link in our unifying and analytical framework for understanding approximate quantum error correction. Another interesting future direction will be to perform the transpose channel recovery on experimental implementations of approximate codes. The transpose channel, like any CPTP map, can certainly be implemented physically using operations on an extended Hilbert space. The more fruitful question, however, will be to discover simpler and more efficient ways of implementing the transpose channel on a specific physical system of our choice.

## V. ACKNOWLEDGMENTS

P.M. would like to thank David Poulin, John Preskill and Todd Brun for useful discussions. H.K.N is supported by the National Research Foundation and the Ministry of Education, Singapore.

## Appendix A: Proof of Corollary 7

Before we can prove Corollary 7, we need the following lemma:

**Lemma 8.** *Consider a subsystem code  $\mathcal{C}$  (of the form given in Eq. (2)), under noise  $\mathcal{E}$ . For any pure state  $\psi_A \equiv |\psi_A\rangle\langle\psi_A|$ ,*

$$1 - \eta_{\text{op}} \left\{ \psi_A \otimes \frac{P_B}{d_B} \right\} \leq \sqrt{[1 + (d_A - 1)\eta_{\text{op}}\{\mathcal{C}\}][1 - \eta_P \left\{ \psi_A \otimes \frac{P_B}{d_B} \right\}]} \quad (\text{A1})$$

The proof of this lemma, which we provide below, follows very closely the proof used to demonstrate Eq. (20) in [1], except that we now have to account for a nontrivial  $B$  subsystem.

*Proof.* Let us denote the Kraus operators of  $\mathcal{R}_{\text{op}}$  for  $\mathcal{C}$  by  $R_i$ . We begin by considering

$$F^2 \left( |\psi_A\rangle, \text{tr}_B \left[ (\mathcal{R}_{\text{op}} \circ \mathcal{E}) \left( \psi_A \otimes \frac{P_B}{d_B} \right) \right] \right) = \frac{1}{d_B} \sum_{i,j,s,t} |\langle \psi_A s_B | R_i E_j | \psi_A t_B \rangle|^2 \equiv \frac{1}{d_B} \sum_{ijst} |Z_{(is)(jt)}|^2,$$

Here,  $\{|s_B\rangle\}$  and  $\{|t_B\rangle\}$  denote orthonormal bases for  $B$ .  $Z$  is a two-index matrix with elements  $Z_{(is)(jt)} \equiv \langle \psi_A s_B | R_i E_j | \psi_A t_B \rangle$ . Invoking the same arguments of diagonalization used to prove our perfect OQEC conditions (see Theorem 1), we have,

$$F^2 \left( |\psi_A\rangle, \text{tr}_B \left[ (\mathcal{R}_{\text{op}} \circ \mathcal{E}) \left( \psi_A \otimes \frac{P_B}{d_B} \right) \right] \right) = \frac{1}{d_B} \sum_{i,s} |Y_{is}^\dagger X_{is}|^2,$$

where  $X_{is} \equiv \mathcal{E}(P)^{-1/4} E_i |\psi_A s_B\rangle$  and  $Y_{is}^\dagger \equiv \langle \psi_A s_B | R_i \mathcal{E}(P)^{1/4}$ . Using the Cauchy-Schwarz inequality twice, we have,

$$F^2 \left( |\psi_A\rangle, \text{tr}_B \left[ (\mathcal{R}_{\text{op}} \circ \mathcal{E}) \left( \psi_A \otimes \frac{P_B}{d_B} \right) \right] \right) \leq \frac{1}{d_B} \sqrt{\sum_{is} |X_{is}^\dagger X_{is}|^2} \sqrt{\sum_{jt} |Y_{jt}^\dagger Y_{jt}|^2}. \quad (\text{A2})$$

To bound the  $X$ -term, consider

$$\begin{aligned} & \sum_{is} |X_{is}^\dagger X_{is}|^2 \\ & \leq \sum_{ijst} |\langle \psi_A t_B | E_j^\dagger \mathcal{E}(P)^{-1/2} E_i | \psi_A s_B \rangle|^2 \\ & = d_B F^2 \left( |\psi_A\rangle, \text{tr}_B \left[ (\mathcal{R}_P \circ \mathcal{E}) \left( \psi_A \otimes \frac{P_B}{d_B} \right) \right] \right). \end{aligned} \quad (\text{A3})$$

To bound the  $Y$ -term, consider

$$\begin{aligned} \sum_{jt} |Y_{jt}^\dagger Y_{jt}|^2 &= \sum_{jt} |\langle \psi_A t_B | R_j^\dagger \mathcal{E}(P)^{1/2} R_j | \psi_A t_B \rangle|^2 \\ &= F^2[|\psi_A\rangle, \mathcal{L}(\psi_A)], \end{aligned} \quad (\text{A4})$$

where  $\mathcal{L}$  is defined as the channel on  $A$  with Kraus operators  $\{\langle t_B | R_j \mathcal{E}(P)^{1/2} R_j^\dagger | t_B \rangle\}$ . Since  $F^2[|\psi_A\rangle, \mathcal{L}(\psi_A)] \leq \text{tr}[\mathcal{L}(\psi_A)]$ , we have,

$$\begin{aligned} \sum_{jt} |Y_{jt}^\dagger Y_{jt}|^2 &\leq \sum_{it} \text{tr} \left[ \langle t_B | R_i \mathcal{E}(P)^{1/2} R_i^\dagger | t_B \rangle | \psi_A \rangle \right. \\ &\quad \left. \times \langle \psi_A | \langle t_B | R_i \mathcal{E}(P)^{1/2} R_i^\dagger | t_B \rangle \right] \\ &\leq \sum_{it} \langle \psi_A | \langle t_B | R_i \mathcal{E}(P) R_i^\dagger | t_B \rangle | \psi_A \rangle \\ &= \langle \psi_A | \text{tr}_B[(\mathcal{R}_{\text{op}} \circ \mathcal{E})(P)] | \psi_A \rangle, \end{aligned} \quad (\text{A5})$$

where in the next-to-last line, we have added in positive terms and used the fact that  $\mathcal{R}_{\text{op}}$  is TP on its domain. Choose an orthonormal basis  $\{|\psi_A^i\rangle\}, i = 1, \dots, d_A$  on  $A$ , such that  $|\psi_A^1\rangle \equiv |\psi_A\rangle$ . Defining  $\alpha_k^i \equiv \langle \psi_A^k | \text{tr}_B[(\mathcal{R}_{\text{op}} \circ \mathcal{E})(|\psi_A^i\rangle\langle\psi_A^i| \otimes P_B/d_B)] | \psi_A^k \rangle$ , we have

$$\begin{aligned} & \sum_{k=1}^{d_A} \langle \psi_A^k | \text{tr}_B[(\mathcal{R}_{\text{op}} \circ \mathcal{E})(|\psi_A^i\rangle\langle\psi_A^i| \otimes P_B/d_B)] | \psi_A^k \rangle \\ &= \sum_{k=1}^{d_A} \alpha_k^i = 1 \quad \forall i = 1, \dots, d_A, \end{aligned} \quad (\text{A6})$$

where the last equality relies on the fact that  $\mathcal{R}_{\text{op}} \circ \mathcal{E}$  is CPTP. Recalling that  $\langle \psi_A | \text{tr}_B[(\mathcal{R}_{\text{op}} \circ \mathcal{E})(\psi_A \otimes P_B/d_B)] | \psi_A \rangle \geq 1 - \eta_{\text{op}}$  for all  $|\psi_A\rangle$ , we have,

$$\begin{aligned} \alpha_i^i &\geq \sum_k \alpha_k^i - \eta_{\text{op}}, \quad \forall i = 1, \dots, d_A \\ \Rightarrow \sum_k \alpha_{k \neq i} &\leq \eta_{\text{op}}, \quad \forall i. \end{aligned} \quad (\text{A7})$$

Going back to Eq. (A5), we can now bound  $\langle \psi_A | \text{tr}_B[(\mathcal{R}_{\text{op}} \circ \mathcal{E})(P)] | \psi_A \rangle$  as follows:

$$\begin{aligned} & \langle \psi_A | \text{tr}_B[(\mathcal{R}_{\text{op}} \circ \mathcal{E})(P)] | \psi_A \rangle \\ &= d_B \sum_i \alpha_1^i = d_B \left( \alpha_1^1 + \sum_{i \neq 1} \alpha_1^i \right) \\ &\leq d_B (1 + (d_A - 1)\eta_{\text{op}}). \end{aligned} \quad (\text{A8})$$

Putting together Eqs. (A2), (A3) and (A8), we obtain our final result Eq. (A1).  $\square$

As a side remark, note that considering  $\mathcal{C}_0$  from Eq. (21) and replacing  $\mathcal{R}_{\text{op}}$  for  $\mathcal{C}$  by  $\mathcal{R}_{\text{op}}$  for  $\mathcal{C}_0$  in the proof of the lemma above provides an alternate and direct proof of Corollary 5.

Now, we are ready to prove Corollary 7:

**Corollary 7** *Given that Condition 1 is satisfied, for a subsystem code  $\mathcal{C}$ ,*

$$\eta_P \leq (d_A + 1)\eta_{\text{op}} + 3\delta + O(\delta^2, \eta_{\text{op}}^2, \delta\eta_{\text{op}}). \quad (\text{A9})$$

*Proof.* For channel  $\mathcal{E}$  satisfying Condition 1, for any recovery  $\mathcal{R}$  and any state  $\psi_A \otimes \rho_B \equiv |\psi_A\rangle\langle\psi_A| \otimes \rho_B$  in  $\mathcal{C}$ ,

$$F^2[|\psi_A\rangle, \text{tr}_B\{(\mathcal{R} \circ \mathcal{E})(\psi_A \otimes \rho_B)\}] \leq \delta + F^2[|\psi\rangle_A, \text{tr}_B\{(\mathcal{R} \circ \mathcal{E})(\psi_A \otimes P_B/d_B)\}], \quad (\text{A10})$$

$$\Rightarrow \eta_{\mathcal{R}}\{\psi_A \otimes \rho_B\} \geq \eta_{\mathcal{R}}\{\psi_A \otimes P_B/d_B\} - \delta. \quad (\text{A11})$$

Interchanging the roles of  $\rho_B$  and  $P_B/d_B$  in Eq. (A10) yields, similarly,

$$\Rightarrow \eta_{\mathcal{R}}\{\psi_A \otimes P_B/d_B\} \geq \eta_{\mathcal{R}}\{\psi_A \otimes \rho_B\} - \delta. \quad (\text{A12})$$

We have the following sequence of inequalities:

$$\begin{aligned} & \eta_{\text{op}}\{C\} \quad (\text{A13}) \\ & \geq \eta_{\text{op}}\{\psi_A \otimes \rho_B\} \\ & \geq \eta_{\text{op}}\{\psi_A \otimes P_B/d_B\} - \delta \quad (\text{using Eq. (A11)}) \\ & \geq 1 - \sqrt{[1 + (d_A - 1)\eta_{\text{op}}\{C\}](1 - \eta_P\{\psi_A \otimes P_B/d_B\})} - \delta \\ & \quad (\text{using Eq. (A1)}) \\ & \geq 1 - \sqrt{[1 + (d_A - 1)\eta_{\text{op}}\{C\}](1 - \eta_P\{\psi_A \otimes \rho_B\} + \delta)} - \delta \\ & \quad (\text{using Eq. (A12)}) \end{aligned}$$

- 
- [1] H. K. Ng and P. Mandayam, Phys. Rev. A **81**, 62342 (2010).
  - [2] A. Ekert and C. Macchiavello, Phys. Rev. Lett. **77**, 2585 (1996).
  - [3] C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters, Phys. Rev. A **54**, 3824 (1996).
  - [4] E. Knill and R. Laflamme, Phys. Rev. A **55**, 900 (1997).
  - [5] D. W. Leung, M. A. Nielsen, I. L. Chuang, and Y. Yamamoto, Phys. Rev. A **56**, 2567 (1997).
  - [6] H. Barnum and E. Knill, J. Math. Phys. **43**, 2097 (2002).
  - [7] J. Tyson, J. Math. Phys. **51**, 092204 (2010).
  - [8] C. Bény and O. Oreshkov, Phys. Rev. Lett. **104**, 120501 (2010).
  - [9] J. Renes, eprint (2010), arXiv:1003.1150v2[quant-ph].
  - [10] N. Yamamoto, S. Hara, and K. Tsumura, Phys. Rev. A **71**, 022322 (2005).
  - [11] M. Reimpell and R. F. Werner, Phys. Rev. Lett. **94**, 080501 (2005).
  - [12] R. L. Kosut, A. Shabani, and D. A. Lidar, Phys. Rev. Lett. **100**, 020502 (2008).
  - [13] A. S. Fletcher, P. W. Shor, and M. Z. Win eprint (2007), arXiv:0710.1052[quant-ph].
  - [14] M. Ohaya and D. Petz, Quantum Entropy and its Use (Springer-Verlag, 1993).
  - [15] A. S. Fletcher, PhD Thesis (MIT), eprint (2007), arXiv:0706.3400[quant-ph].
  - [16] D.W. Kribs, R. Laflamme and D. Poulin, Phys. Rev. Lett. **94**, 180501 (2005).
  - [17] D.W. Kribs, R. Laflamme, D. Poulin and M. Lesosky, Quant. Inf. Comp. **6**, 382 (2006).
  - [18] M. A. Nielsen and D. Poulin, Phys. Rev. A **75**, 064304 (2007).
  - [19] D. Poulin, Phys. Rev. Lett. **95**, 230504 (2005).
  - [20] D. Bacon, Phys. Rev. A **73**, 012340 (2006).
  - [21] P. Aliferis and A. Cross, Phys. Rev. Lett. **98**, 220502 (2007).
  - [22] S. Aly, A. Klappenecker, and P. Sarvepalli, eprint (2006), arXiv:quant-ph/0610153.
  - [23] D. Petz, Rev. Math. Phys. **15**, 79 (2003).
  - [24] P. Hayden, R. Jozsa, D. Petz and A. Winter, Comm. Math. Phys. **246**, 359 (2004).
  - [25] R. Blume-Kohout, H. K. Ng, D. Poulin, and L. Viola, Phys. Rev. Lett. **100**, 030501 (2008).
  - [26] R. Blume-Kohout, H. K. Ng, D. Poulin, and L. Viola, Phys. Rev. A **82**, 062306 (2010).
  - [27] A. Shabani and D. Lidar, Phys. Rev. A **72**, 042303 (2005).
  - [28] Instead of considering  $\mathcal{C}$  to consist only of *product* states on  $AB$ , another interesting possibility is to choose  $\mathcal{C}$  to be *all* states on  $AB$ , including all separable and entangled states. However, in this case, it is not so clear where the information initially resides. Once there is correlation between subsystems  $A$  and  $B$ , be it classical or quantum, one can no longer say that the information resides only in subsystem  $A$ . One can perhaps say that, for an initial state  $\rho$  on  $AB$ , the information lies in subsystem  $A$  represented by  $\rho_A = \text{tr}_B(\rho)$ . However, if we suppose that we have complete control over preparation of  $A$  and  $B$ , then we might as well have prepared the state  $\rho_A \otimes \rho_B$  instead of  $\rho$ , without losing any information. Hence, we will restrict ourselves to codes consisting only of product states. The alternative case will be of interest for future studies where experimental difficulties result in the inability to prepare only product initial states in  $AB$  (see,



for example, [27] for noiseless subsystems without a priori initialization into product states).

[29] In practice, it is useful to use a decomposition Eq. (1) that is not arbitrarily invented by the experimenter, but that is induced by the structure of the noise afflicting the quantum system, so that one can identify a subsystem

in the decomposition that best ensures longevity of the stored information.

[30] Note that  $B_{ij}$ , in this case of subspace codes, is just a scalar and  $\Delta_{ij}$  is an operator on  $A$  only.